Pergamon

Topology Vol. 35, No. 3, pp. 779–806, 1996

0040-9383/96 $15.00 + 0.00

0040-9383(95)00035-6

ON THE CONJUGACY PROBLEM FOR AUTOMORPHISMS OF FREE GROUPS

JÉRÔME E. LOS

(*Received in final form* 16 *May* 1994)

1. INTRODUCTION

THE CONJUGACY problem is one of the basic group theoretic decision problems which has been formulated and solved in certain cases by Dehn in 1912 [4]. For quite a large class of groups an answer is known about the existence or the non-existence of a solution to this problem (see [10] for a review article). In this paper we study $Out(F_n)$, the group of automorphisms of a free group of rank n, modulo inner automorphisms. An outer automorphism Φ is called *reducible* if there are proper free factors $F_1, \ldots, F_k$ of F_n such that Φ transitively permutes the conjugacy classes of the F_i's and such that $F_1 * F_2 * \ldots * F_k$ is a free factor of F_n. If Φ is not reducible it is called *irreducible*. The main result of this paper is:

THEOREM 1. *The conjugacy problem for the irreducible outer automorphisms of a free group admits a solution.*

It is well known that $Out(F_n)$ has many similarities with the mapping class group $MCG(S)$ of a compact surface. For this group a solution of the conjugacy problem has been found by Hemion [9] and later on by Mosher [14]. In this case, according to the Nielsen–Thurston classification theorem [15, 17], the conjugacy problem can be reduced to the pseudo-Anosov case, which is the class analogous to the non-periodic irreducible elements of $Out(F_n)$. Unfortunately, for $Out(F_n)$ the irreducible elements are not the whole story and our Theorem 1 is not sufficient to solve the conjugacy problem in full generality.

Let us now describe the strategy of the proof. A *marked graph* G is a graph whose fundamental group is identified with F_n. The set of all marked graphs plus some additional metric structures has been called the outer space by Culler and Vogtmann [3]. Each element Φ of $Out(F_n)$ can be represented non-uniquely by a homotopy equivalence $f: G \to G$ on a marked graph. We can assume further that f maps the set of vertices to itself (not necessarily one to one) and that the restriction of f to each edge is locally injective. Such a map is called a *topological representative* of Φ. Among the topological representatives some of them have much better properties, namely the train track maps as defined by Thurston and Bestvina and Handel [2]. A topological representative $f: G \to G$ is a *train track* map if the restriction of f^k to every edge of G is locally injective for all $k > 0$.

The goal of this paper is to study some structures of the set of train track maps $\mathcal{T}(\Phi)$ for a given automorphism Φ. An easy property is that $\mathcal{T}(\Phi)$ is finite; we prove further that it is a complete conjugacy invariant, see Section 4. In order to solve the conjugacy problem we have to be able to describe the whole set $\mathcal{T}(\Phi)$. Such a description is given in Section 3; it is formulated as a path connected property.

The starting point is a theorem due to Bestvina and Handel [2] stating that every irreducible automorphism Φ of F_n admits a train track representative which can be reached by a finite algorithm. This algorithm is improved in Section 2 by simplifying the combinatorics of the algorithm. This is obtained by reducing the number of operations from 6 to 4, namely we avoid the valence 1 and 2 isotopies.

Among the irreducible elements of $Out(F_n)$ the periodic ones are much simpler than the others from many points of view. This (finite) subclass has been studied by Dicks and Ventura [6] and their results include a solution to the conjugacy problem in this special class. We can therefore restrict our study to the non-periodic irreducible elements which is the "generic" case.

Notice that the train track maps, in the same context of free group automorphisms, have been defined in a slightly different way by Lustig [13]. For surface homeomorphisms some closely related algorithms appeared recently in [1, 11], the goal being to decide, among other properties, whether a mapping class is pseudo-Anosov, reducible or finite order. These results also enable to describe the invariant foliations, to find a Markov partition, to compute the dilatation factor, etc.

The author is glad to acknowledge M. Boileau, W. Dicks, M. Handel, G. Levitt, F. Paulin and R. Penner for discussions on this question which have been very helpful at different stages of this work.

2. TRAIN TRACK MAPS: AN IMPROVED ALGORITHM

2.1. Basic properties

In this section we recall some basic results from [2] and we give an improved version of the algorithm. This new version avoids several of the operations used in the previous algorithm such as the valence 1 and 2 isotopies. The combinatorics of the algorithm is then much simpler. The background definitions on graphs and maps on graphs can be found in the papers by Stallings or Gersten [16, 8].

As for every algorithm, one needs an initial condition. The initial data being an automorphism Φ of a free group F_n, we first "geometrize" these data in a standard way. Let R_n be a rose with n petals, i.e. a graph with n edges and one vertex, denoted by v_0. Let us fix an identification between F_n and $\pi_1(R_n, v_0)$, then we identify $\Phi: F_n \to F_n$ with $f_\#: \pi_1(R_n, v_0) \to \pi_1(R_n, v_0)$ which is induced by a homotopy equivalence $f: R_n \to R_n$.

A *marked graph* is a graph whose fundamental group, with respect to a vertex, has been identified with F_n. In other words a marked graph G is a graph along with a homotopy equivalence $R_n \to G$. If $f: G \to G$ is a homotopy equivalence it defines an outer automorphism of $\pi_1(G, v)$, where v is any vertex of G. Let us denote by $E(G)$ and $V(G)$ respectively the sets of edges and vertices of the graph G. Each edge of the graph, $e \in E(G)$, is oriented and connects its *initial vertex* $i(e)$ to its *terminal vertex* $t(e)$. Let us denote by $\bar{e}$ the same edge but with the reverse orientation $(i(\bar{e}) = t(e))$. If G is a marked graph with $E(G) = \{e_1, \dots, e_k\}$, then an *edge path* in G is a word: $p = x_1.x_2 \dots x_r$, where $x_i \in E(G) \cup \overline{E(G)}$ and with the restriction that $t(x_j) = i(x_{j+1})$ for all $j = 1, \dots, r-1$. The initial vertex of the edge path p is $i(p) = i(x_1)$ and its terminal vertex is $t(p) = t(x_r)$. If an edge path has no backtracking (or cancellation), i.e. a subword of the form $x.\bar{x}$, then it is called *reduced*.

If G is a marked graph and $E(G) = \{e_1, \dots, e_k\}$ then a homotopy equivalence $f: G \to G$ is given by a k-tuple of edge paths $\{f(e_1), \dots, f(e_k)\}$. By definition, f is a *topological*

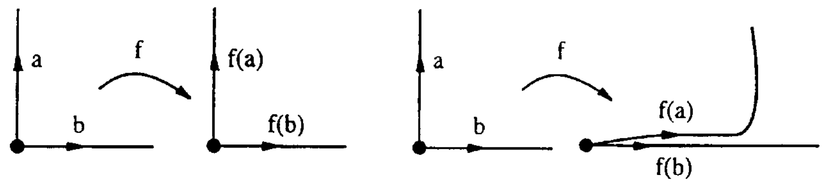

Fig. 1. Transversality-tangency.

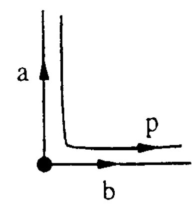

Fig. 2. A path crossing a turn.

representative of $\Phi \in Out(F_n)$ if every $f(e_i)$ is reduced (locally injective). Let us denote by $\mathscr{R}_n(\Phi)$ the set of all the topological representatives of the automorphism Φ. The *transition matrix* $M(f, G)$ of the topological representative $f: G \to G$ is the $k \times k$ integer matrix whose entry $m_{i,j}$ is the number of times the letter e_i (or $\overline{e_i}$) occurs in the edge path $f(e_j)$. The topological representative $f: G \to G$, which is denoted as the pair (f, G), is *irreducible* if the matrix $M(f, G)$ is irreducible. The *growth rate* (or dilatation) $\lambda(f, G)$ is the largest eigenvalue of the transition matrix $M(f, G)$.

A proper subgraph (not reduced to a vertex) G_0 of G is *invariant* under $f: G \to G$ if $G_0 \supset f(G_0)$. In this case the transition matrix is reducible (i.e. there is a block decomposition) and, conversely, the same holds true. A *tree* is a connected contractible graph (not reduced to a vertex) and a *forest* is a graph whose connected components are trees. An outer *automorphism* Φ *is irreducible* if every topological representative without invariant forest $(f, G) \in \mathscr{R}_n(\Phi)$ is irreducible. This definition is equivalent to the one given in the introduction (see [2]). Notice that we assumed implicitly that the graphs have no valence one vertices.

Let us introduce some more terminology. For a given vertex $v \in V(G)$ the *star at* v is the set of edges which are incident at v, i.e. $St(v) = \{e \in E(G) | v = i(e) \text{ or } v = t(e)\}$. The *valence* $Val(v)$ of a vertex v is the cardinality of $St(v)$. A *turn* $(a^{\varepsilon_a}, b^{\varepsilon_b})$ at $v \in V(G)$ is an unordered pair of distinct oriented edges in $St(v)$, where the notation means $a^{\varepsilon_a} = a$ if $v = i(a)$ and $a^{\varepsilon_a} = \bar{a}$ if $v = t(a)$. We say that the two edges of a turn are *transverse* under $f: G \to G$ if the f-image of a turn is also a turn. Conversely, if the f-image of a turn is not a turn then the two edges are called *tangent* under f. To be more specific let (a, b) be a turn; the f-image of this turn is transverse if the two edge paths $f(a)$ and $f(b)$ start with distinct letters. Conversely, if the two edge paths start with the same letter then the two edges are tangent under f (see Fig. 1).

An edge path p in G is said to *cross the turn* (a, b) if the subpath $\bar{a} \cdot b$ or $\bar{b} \cdot a$ occurs in p (see Fig. 2).

Assume that the f-image of an edge $e \in E(G)$ crosses the turn (a, b); for instance $f(e) = X \cdot \bar{a} \cdot b \cdot Y$, where X and Y are some edge paths compatible with this notation. If, in addition, the turn (a, b) is tangent under f, i.e. $f(a) = \alpha \cdot M$ and $f(b) = \alpha \cdot N$, then the second iterate of the edge e is non-locally injective since $f^2(e) = f(X) \cdot \bar{M} \cdot \bar{\alpha} \cdot \alpha \cdot N \cdot f(Y)$. The goal of the algorithm is to "remove" all these situations which create non-local injectivity (cancellations).

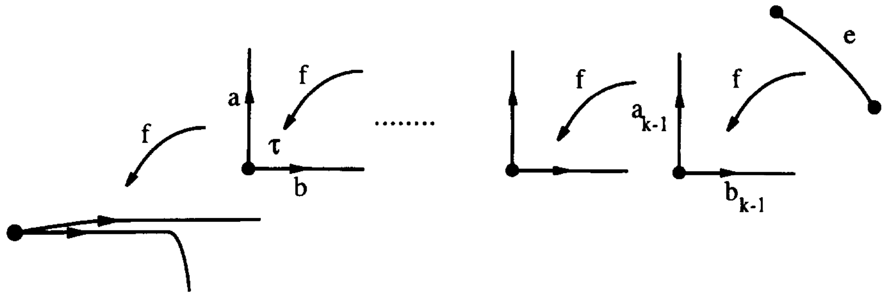

Fig. 3. The case $c(\tau) = k > 1$.

Definition 2.1.1. Let (f, G) be a topological representative of $\Phi \in Out(F_n)$ and let τ be a turn in G. The *tangent coefficient* $c(\tau, f) \in N$ is the smallest integer k so that there exists an edge $e \in E(G)$ such that $f^k(e)$ crosses τ. If no such edge exists then we set $c(\tau, f) = \infty$. If no confusion is possible then we shall omit to specify the map f.

We described above the case of a tangent turn τ with a tangent coefficient $c(\tau, f) = 1$. Let us now describe the case $c(\tau, f) = k > 1$. This situation occurs if and only if there is an edge e whose f-image crosses a transverse turn (a_{k-1}, b_{k-1}) which is mapped under f to a transverse turn (a_{k-2}, b_{k-2}) and so on until the transverse turn (a_1, b_1) is mapped to the tangent turn τ (see Fig. 3). In other words, if there exists a tangent turn τ such that $c(\tau, f) = k \neq \infty$ then the $(k+1)$st iterate of the map f admits a cancellation at this turn.

2.2. *The moves*

Let us now describe the elementary moves which enable one to transform the initial topological representative (f, R_n) to a train track map.

(1) *Collapsing an invariant forest*: Assume that $f: G \to G$ has an invariant forest G_0 such that $E(G_0) = \{e_1, \ldots, e_r\}$. We change the graph and the map by collapsing each edge of G_0. At the edge path level this transformation is given by

(1a) removing all the images $f(e_i)$ for all $i = 1, \ldots, r$ and

(1b) removing all occurrences of the letters $\{e_1^{\pm}, \ldots, e_r^{\pm}\}$ from the words $f(e_j)$ for all $j > r$.

(2) *Collapsing a pretrivial forest*: Assume that G_0 is a forest such that some iterate $f^m(G_0)$ is a collection of vertices. Such a forest is called a *pretrivial*. Collapsing G_0 defines a new topological representative $f': G' \to G'$.

(3) *Pulling tight*: Assume that $f: G \to G$ is not a topological representative because some edge paths $f(e_i)$ are non-locally injective. These non-local injectivities come from cancellations in the words. The new topological representative $f': G \to G$ is obtained from f by removing these cancellations.

(4) *Folding*: This operation is surely the most important. It has been introduced by Dicks [5] and by Stallings [16] and has been used as a main tool in [2]. Here we introduce a slight improvement of the operation. The goal is to remove the possible cancellations by suppressing all the "dangerous" tangent turns, i.e. the tangent turns having finite tangent coefficient.

Let us assume that the turn $\tau = (a, b)$ is tangent under $f: G \to G$ and that $c(\tau) < \infty$. The tangency implies, with the given orientations, that the two edge paths $f(a)$ and $f(b)$ have a common initial edge path, i.e.

$$f(a) = M \cdot A \quad \text{and} \quad f(b) = M \cdot B,$$

where M is the longest common edge path.

(4a) We first transform the graph G to the graph G' by identifying the beginning of the edges a and b (with the chosen orientation). This operation on the graph is called a *splitting*. Notice that the topology of the graph does not change if the vertex v at which the edges a and b start has valence three. If $Val(v) > 3$ then this operation creates a new edge $\{x\}$ and a new vertex $\{w\}$.

(4b) We then transform the image from $f: G \to G$ to $f': G' \to G'$ in the following way (with the previous notations):

$$E(G) = \{a, b, e_3, \ldots, e_N\} \to E(G') = \{a', b', x, e'_3, \ldots, e'_N\}$$

$$f'(a') = (A)^*, \quad f'(b') = (B)^*, \quad f'(x) = (M)^*$$

$$f'(e'_j) = (f(e_j))^* \quad \text{for all } e_j \neq a, b, x.$$

The transformation $X \to (X)^*$ is a rewriting of an edge path X in G to an edge path $(X)^*$ in G'. This transformation is given by

$$a \to xa', \quad b \to xb' \quad \text{and} \quad e_j \to e'_j \text{ for all other edges } e_j \neq a, b. \tag{$*$}$$

This rewriting depends upon the chosen orientation of the three edges a, b and x. The adaptation to the other cases is obvious.

Several particular cases have to be considered.

(a) If the edge paths A and B are both non-trivial (non-reduced to a vertex) then the operation is called a *partial folding*. The new vertex has valence 3 in this case.

(b) If either A or B is trivial then the respective image $f'(a')$ or $f'(b')$ is a vertex and then we can collapse the pretrivial tree which is either a' or b'. Such a folding is called *absorbing* (or full).

(c) If the vertex v at which the tangent turn occurs has valence 3, we observed above that the topology of the graph has not changed. Let us call $\{y\}$ the third edge of $St(v)$ in G, for instance with $t(y) = v$. Formally the folding operation creates a new edge $\{x\}$ which connects v' to v'', with $Val(v') = 2$ and $Val(v'') \geqslant 3$, i.e. such that $St(v') = \{y, x\}$ and $St(v'') = \{a', b', x, \ldots\}$. We want to avoid considering valence two vertices. To that end we identify the new edge $\{x\}$ with the end of edge $\{y\}$. This is done in two steps; first we change the above transformation $(*)$ in the following way. An edge path X in G is transformed into an edge path X^{**} in G by

$$e_j \to e'_j \quad \text{for all the edges (including } a \text{ and } b\text{);} \tag{$**$}$$

in other words the transformation $(**)$ is the identity. Then we transform the image $f(y)$ (depending upon the chosen orientation) into

$$f'(y) = (f(y))^{**} \cdot (M)^{**};$$

the other edges are transformed, as above, according to

$$f'(a') = (A)^{**}, \quad f'(b') = (B)^{**}$$

and

$$f'(e'_j) = (f(e_j))^{**} \quad \text{for all } e_j \neq a, b, y.$$

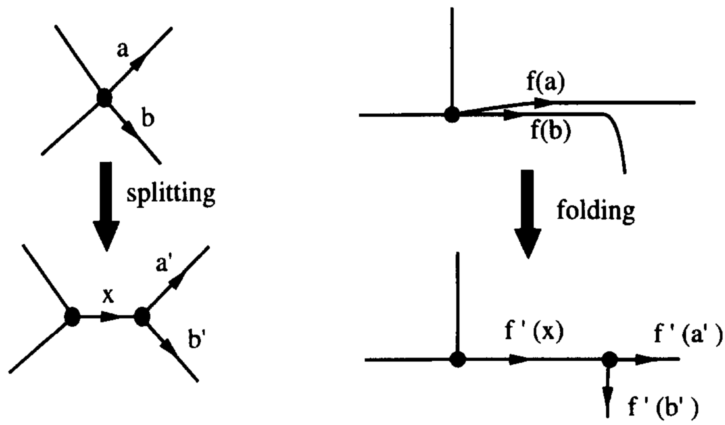

Fig. 4. Folding.

This last transformation is, in the terminology of [2], a folding followed by a valence two isotopy.

Let us introduce some notations before going on. We denote by $T(f, G)$ the set of all the tangent turns of the topological representative (f, G) and by

$$m(f, G) = \inf\{c(\tau_i) | \tau_i \in T(f, G)\}.$$

Let us now describe the effects of these elementary operations on the growth rate $\lambda(f, G)$.

(2.2.1) After collapsing a forest, we obtain (f', G') such that $\lambda(f', G') \leqslant \lambda(f, G)$.
(2.2.2) After pulling tight the growth rate goes down: $\lambda(f', G') < \lambda(f, G)$.

The proof of these two properties is an easy matrix computation and can be found in [2]. For the folding operation, several cases are possible.

LEMMA 2.2.3. *Assume that a topological representative $(f, G) \in \mathscr{R}_n(\Phi)$ has a tangent turn τ at a valence k vertex. Then a folding operation at τ defines a map $f': G' \to G'$ so that:*

(1) *If $k > 3$ and $f': G' \to G'$ is a topological representative then $\lambda(f', G') = \lambda(f, G)$.*

(2) *If $k > 3$ and $f': G' \to G'$ is not a topological representative then, after pulling tight and eventually collapsing a pretrivial or an invariant forest, we obtain a topological representative (f'', G'') such that $\lambda(f'', G'') < \lambda(f, G)$.*

(3) *If $k = 3$ and if $c(\tau) = m(f, G)$ then the folding operation defines a topological representative (f', G') such that*

(i) *$\lambda(f', G') < \lambda(f, G)$ if $1 \leqslant c(\tau) < \infty$, and*
(ii) *$\lambda(f', G') = \lambda(f, G)$ if $c(\tau) = \infty$.*

The first two statements are just a reformulation of Lemma 1.15 in [2]. Let us observe that if $k > 3$ and $c(\tau) > 1$ then we are necessarily in case (1) of the lemma. It remains to prove part (3) of the lemma.

Let us first consider the case $k = 3$ and $c(\tau) = 1$. As a first step we prove that $\lambda(f', G') \leqslant \lambda(f, G)$. The fact that the two graphs G and G' are the same and that the rewriting operation (**) defined above is the identity enables one to compare directly the two incidence matrices $M(f, G)$ and $M(f', G')$. In the vector space of $k \times k$ matrices we consider the norm: $\|M\| = \sum_{i,j}|m_{i,j}|$. The folding operation defined above implies that

$$\|M(f', G')\| = \|M(f, G)\| - l(M) \tag{1}$$

where l is the geometric length of the common edge path M. Indeed, the changes between (f, G) and (f', G') occur only at the edge paths $f(a), f(b)$ and $f(y)$ as

$$l(f'(a')) = l(f(a)) - l(M)$$

$$l(f'(b')) = l(f(b)) - l(M)$$

$$l(f'(y)) = l(f(y)) + l(M)$$

which prove relation (1).

We can also compare the norms of the powers of $M(f, G)$ and $M(f', G')$. Indeed, since $c(\tau) = 1$ then $f^{(2)}: G \to G$ admits a cancellation at the edge e such that $f(e)$ crosses the turn (a, b). The length of this cancellation is $2l(M)$. The folding operation defined above implies that this cancellation has been removed in (f', G') since the tangency has been removed. This implies that

$$\|M^2(f, G)\| - \|M^2(f', G')\| \geqslant l(f(M)) + 2l(M). \tag{2}$$

This argument, used inductively, proves that for every integer n we have

$$\|M^n(f, G)\| - \|M^n(f', G')\| \geqslant l(f^{(n-1)}(M)) + 2l(f^{(n-2)}(M)). \tag{3}$$

Let us recall that the largest eigenvalue of the matrix M can be computed by

$$\lambda(M) = \lim_{n \to \infty} \sqrt[n]{\|M^n\|}.$$

Therefore inequality (3) implies

$$\lambda(f', G') \leqslant \lambda(f, G). \tag{4}$$

Let us observe that the assumption $c(\tau) = 1$ has been used to prove (2). If $c(\tau) = k > 1$ then the same arguments show that inequality (3) is valid up to a shift in the power of the last term because the cancellation occurs for the $(k + 1)$st iterate of f and has been removed for f'. Therefore inequality (4) is also valid in this case.

Let us now prove that inequality (4) is actually strict. As above we first consider the case $c(\tau) = 1$. Let $\mu(f, G)$ be the positive eigenvector of $M(f, G)$ for the eigenvalue $\lambda(f, G)$. From the Perron–Frobenius theorem (see for instance [7]), inequality (4) above implies

$$\mu_a(f, G) + \mu_b(f, G) \geqslant \mu_y(f, G). \tag{5}$$

Let us now consider the second iterate of each map $f^{(2)}: G \to G$ and $f'^{(2)}: G' \to G'$. The incidence matrices of these maps are, respectively, $M^{(2)}(f, G)$ and $M^{(2)}(f', G')$. The differences between these two maps concern now the edges $e_i \in E_M \subset E(G)$ (resp. $E_{f(M)} \subset E(G)$) which are contained in the edge paths M (resp. $f(M)$). Recall that the cancellation of $f^{(2)}$ at the edges $e_i \in E_M$ has been removed for $f'^{(2)}$. Therefore inequality (5) implies, since $\mu_{e_i}(f, G) > 0$, $\forall e_i \in E(G)$, that

$$(M^2(f, G)\cdot\mu)_i \geqslant (M^2(f', G')\cdot\mu)_i, \quad \forall i \text{ s.t. } e_i \notin E_M$$

$$(M^2(f, G)\cdot\mu)_j > (M^2(f', G')\cdot\mu)_j, \quad \forall j \text{ s.t. } e_j \in E_M.$$

We conclude, using the Perron–Frobenius theorem, that

$$\lambda(f', G') < \lambda(f, G). \tag{6}$$

It remains to consider the case $c(\tau) = k > 1$.

We assumed that $c(\tau) = \inf\{c(\tau_i);\ \tau_i$ is a tangent turn of $(f, G)\}$. From the definition of the tangent coefficient, the map $f^{(k)}: G \to G$ is a topological representative which admits a tangent turn τ whose tangent coefficient is 1. Therefore the previous proof implies that $\lambda^k(f', G') < \lambda^k(f, G)$, which completes the proof.

In the case $c(\tau) = \infty$, from the definition of the tangent coefficient and of a train track representative one has:

PROPOSITION 2.2.4. *An irreducible topological representative (f, G) is a train track representative if and only if $m(f, G) = \infty$.*

Therefore the last case (3)(ii) of Lemma 2.2.3 is obvious.

Let us now consider more precisely the situation where a tangent turn $\tau = (a, b)$ at a vertex v is such that $c(\tau) = m$. As we have seen above, this situation occurs if the following conditions are satisfied:

(T1) $f(a) = M.A$, and $f(b) = M.B$, where M is a non-trivial edge path ($l(M) \neq 0$). The initial vertex of M is $f(v)$ and we denote by $w(\tau, f)$ its terminal vertex.

(T2) There exists an ordered sequence of turns $\{\tau_1, \tau_2, \ldots, \tau_{m-1}\}$ at the vertices $\{v_1, v_2, \ldots, v_{m-1}\}$ such that $\tau_i = (a_i, b_i)$ and τ_i is mapped to τ_{i-1} under f, which means that $f(a_i) = a_{i-1}.\ A_i$ and $f(b_i) = b_{i-1}.\ B_i$ for all $i \in \{1, \ldots, m-1\}$, where τ_0 is identified with τ.

(T3) There exists an edge $e \in E(G)$ such that $f(e)$ crosses the turn τ_{m-1} and no edge path $f(e_i)$ crosses any of the turns $\{\tau, \tau_1, \tau_2, \ldots, \tau_{m-2}\}$. By definition, the set $E(\tau, m) = \{e_i \in E(G) \mid f^m(e_i)$ crosses $\tau\}$ is non-empty.

LEMMA 2.2.5. *With the above notations let $\tau = (a, b)$ be a turn at a vertex v of valency $k > 3$ and such that τ is tangent under f with $c(\tau) = m = m(f, G)$. If either*

(a) *$E(\tau, m) \neq \{a\}$ (resp. $\{b\}$), or*
(b) *$E(\tau, m) = \{a\}$ (resp. $\{b\}$) and $w(\tau, f) \neq v_{m-1}$,*

then we have:

(1) *If $c(\tau) = m(f, G) = 1$, then after the folding operation we are in situation* (2) *of Lemma* 2.2.3 *and thus the growth rate goes down.*

(2) *If $c(\tau) = m(f, G) > 1$, then after the folding operation we obtain a topological representative (f', G') such that $m(f', G') \leqslant m(f, G) - 1$.*

Let us first assume that $c(\tau) = 1$. In case (a) there is an edge $e \neq a, b$ in $E(\tau, m)$ and we can choose the orientation of e in such a way that $f(e) = X \cdot \bar{a} \cdot b \cdot Y$. After the folding operation we obtain a map $f': G' \to G'$ such that, with the previous notations,

$$f'(e') = X^* \cdot \bar{a}' \cdot \bar{x} \cdot x \cdot b' \cdot Y^*$$

and part (2) of Lemma 2.2.3 applies.

In case (b), for instance with $E(\tau, m) = \{a\}$, we have by (T1) that $f(a) = M \cdot A$. The second assumption, i.e. $w(\tau, f) \neq v_{m-1}$, implies that the subpath $\bar{a} \cdot b$ belongs either to the path M or the path A. Therefore the previous argument applies, after the folding operation, either to the new edge x or to the edge a'. This completes the proof of (1).

If $c(\tau) > 1$ then, after the folding operation, we obtain a topological representative (f', G'). The sequence of turns $\{\tau_1, \tau_2, \dots, \tau_{m-1}\}$ is transformed into a sequence $(\tau'_1, \tau'_2, \dots, \tau'_{m-1}\}$ such that τ'_i is mapped to τ'_{i-1} under f'. Moreover, since we have $f(a_1) = a \cdot A_1$ and $f(b_1) = b \cdot B_1$ then, after the folding operation, one gets

$$f'(a'_1) = x \cdot a' \cdot A_1^* \quad \text{and} \quad f'(b'_1) = x \cdot b' \cdot B_1^*$$

and thus the turn $\tau'_1 = (a'_1, b'_1)$ is tangent under f'. In case (a) there is an edge e' in $E(G') - \{a', b', x\}$ such that $f'(e')$ crosses τ'_{m-1}. In case (b) the same is true for either $f'(a')$ or $f'(x)$; therefore $c(\tau'_1, f') = m - 1$ and thus $m(f', G') \leqslant m(f, G) - 1$.

The remaining case is when $E(\tau, m) = \{a\}$ (resp. $\{b\}$) and $w(\tau, f) = v_{m-1}$. With the previous notations this means that:

(i) $f(a) = Q \cdot \bar{b}_{m-1} \cdot a_{m-1} R$ and $f(b) = Q \cdot \bar{b}_{m-1} \cdot B$, where $M = Q \cdot \bar{b}_{m-1}$, $A = a_{m-1} R$ or
(ii) $f(a) = N \cdot \bar{a}_{m-1} \cdot b_{m-1} P$ and $f(b) = N \cdot \bar{a}_{m-1} \cdot B$, where $M = N \cdot \bar{a}_{m-1}$, $A = b_{m-1} P$,
and we set

$$\varepsilon(\tau, f) = \begin{cases} 1 & \text{if case (i) occurs} \\ 2 & \text{if case (ii) occurs} \\ \infty & \text{otherwise.} \end{cases}$$

If, in the ordered sequence $\{v_0 = v, v_1, v_2, \dots, v_{m-1}\}$, some vertices have valence 3, then we denote by $\delta(\tau, f)$ the first integer j such that $Val(v_j) = 3$ and we set $\delta(\tau, f) = \infty$ otherwise. Recall also that the sequence of turns is characterized by the ordered collection of edge paths $\{(M, A, B); (A_1, B_1); \dots; (A_{m-1}, B_{m-1})\}$. Finally we introduce an ordered sequence of integers as follows:

$$S_\varepsilon(\tau, f) = \begin{cases} \{l(B), l(B_1), \dots, l(B_{m-1}), l(M) - 1\} & \text{if } \varepsilon(\tau, f) = 1 \\ \{l(B), l(B_1), \dots, l(B_{m-1}), l(A) - 1, l(A_1), \dots, l(A_{m-1}), l(M) - 1\} & \text{if } \varepsilon(\tau, f) = 2 \end{cases}$$

where $l(..)$ is the length of an edge path.

Now we set $\Delta(\tau, f) = 0$ if $\varepsilon(\tau, f) = \infty$. If $\varepsilon(\tau, f) \neq \infty$ then we set $\Delta(\tau, f) = \infty$ if all the entries of the sequence $S_\varepsilon(\tau, f)$ are zero. Otherwise $\Delta(\tau, f)$ is the position (in the ordered sequence $S_\varepsilon(\tau, f)$) of the first non-zero entry.

Proposition 2.2.6. *With the above notations, if $\Delta(\tau, f) = \infty$ then the topological representative (f, G) of the irreducible outer automorphism Φ either admits an invariant forest or Φ is finite order.*

This result is obvious once the assumption $\Delta(\tau, f) = \infty$ is made explicit. Indeed it means that all the elements of the sequence $S_\varepsilon(\tau, f)$ are zero. Two cases are possible depending upon the value of ε. In the case $\varepsilon = 2$ then one has

$$f(a) = \bar{a}_{m-1} b_{m-1}, \quad f(b) = \bar{a}_{m-1}$$

and

$$f(a_1) = a, \quad f(b_1) = b, \quad f(a_i) = a_{i-1}, \quad f(b_i) = b_{i-1} \qquad \text{for all } i \in \{2, \dots, m-1\}.$$

Therefore the (sub)graph G_0 whose edge set is $E(G_0) = \{a, b, a_1, b_1, \dots, a_{m-1}, b_{m-1}\}$ is invariant under f. If G_0 is a proper subgraph then it is an invariant forest since the automorphism Φ is irreducible.

Otherwise, i.e. if $G_0 = G$, then it is easy to check that the automorphism Φ is finite order. Indeed we check that $f^m(a) = \bar{a}\cdot b$, and $f^m(b) = \bar{a}$ and then that

$$f^{3m}(a) = a \quad \text{and} \quad f^{3m}(b) = b.$$

The other case, $\varepsilon = 1$, is even easier to check and is left to the reader.

Let us now define several other invariants related to a topological representative (f, G). We have already defined $T(f, G)$ as the set of all tangent turns of (f, G) and $m(f, G)$ as $m(f, G) = \min\{c(\tau_i) | \tau_i \in T(f, G)\}$ and we set

$$T_m(f, G) = \{\tau_i \in T(f, G) | c(\tau_i) = m(f, G)\}.$$

For each element $\tau_i \in T_m(f, G)$, the integer $\delta(\tau_i, f)$ is well defined and then we can consider the integer

$$d(f, G) = \min\{\delta(\tau_i, f) | \tau_i \in T_m(f, G)\}$$

together with the set

$$T_d(f, G) = \{\tau_i \in T_m(f, G) | \delta(\tau_i, f) = d(f, G)\}.$$

Similarly, for each $\tau_i \in T_d(f, G)$ the number $\Delta(\tau_i, f)$ is well defined and we consider

$$D(f, G) = \min\{\Delta(\tau_i, f) | \tau_i \in T_d(f, G)\}$$

together with the set

$$T_D(f, G) = \{\tau_i \in T_d(f, G) | \Delta(\tau_i, f) = D(f, G)\}.$$

Definition 2.2.7 Let (f, G) be a topological representative of an irreducible element $\Phi \in Out(F_n)$. We define the following *complexity function*:

$$\Lambda(f, G) = \{\lambda(f, G); m(f, G); d(f, G); D(f, G); e(G)\}$$

where $m(f, G)$, $d(f, G)$, $D(f, G)$ are defined above and $e(G) = Card\, E(G)$, with the ordering relation $\Lambda(f', G') < \Lambda(f, G)$ if

(i) $\lambda(f', G') < \lambda(f, G)$ or
(ii) $\lambda(f', G') = \lambda(f, G)$ and $m(f', G') < m(f, G)$ or
(iii) $\lambda(f', G') = \lambda(f, G)$, $m(f', G') = m(f, G)$ and $d(f', G') < d(f, G)$ or
(iv) $\lambda(f', G') = \lambda(f, G)$, $m(f', G') = m(f, G)$, $d(f', G') = d(f, G)$ and $D(f', G') < D(f, G)$ or
(v) $\lambda(f', G') = \lambda(f, G)$ and $e(G') < e(G)$.

We recall (see [2]) that the growth rate of a train track representative is minimal, among the growth rates of all the topological representatives. Now we can state the main result of this section:

THEOREM 2.2.8. *Let ϕ be a non-periodic irreducible outer automorphism of F_n and let (f, G) be any topological representative of Φ. Then there is a sequence of the elementary operations* (1)–(4) *which defines a sequence of topological representatives:*

$$(f, G) = (f_0, G_0) \to (f_1, G_1) \to \cdots \to (f_K, G_K)$$

whose complexity function $\Lambda(f_i, G_i)$ is strictly decreasing.

It is clear that Theorem 2.2.8, together with the minimality property of the growth rate, implies the existence of a finite sequence of elementary operations ending with a train track representative for the corresponding automorphism.

2.3. Proof of Theorem 2.2.8

The proof of this theorem is based on the argument used in the proof of Lemma 2.2.5. Let (f, G) be any topological representative. If $m(f, G) = \infty$ then, by Lemma 2.2.4, (f, G) is a train track representative and thus we are done.

If $m(f, G) < \infty$ and $d(f, G) = 0$ then there is a tangent turn $\tau \in T_m(f, G)$ at a valency 3 vertex v. Therefore, after a folding operation at τ, we obtain a topological representative (f', G') such that $\lambda(f', G') < \lambda(f, G)$, by Lemma 2.2.3.

In all the cases where Lemma 2.2.5 applies (almost all the cases), i.e. if $\Delta(f, G) = 0$, then either $\lambda(f, G)$ or $m(f, G)$ decreases after one folding operation and thus Λ is decreasing. If (f, G) admits an invariant forest then, after collapsing the forest, we get $e(G') < e(G)$ and the function Λ is decreasing.

In the case where Lemma 2.2.5 does not apply, i.e. if there exists a tangent turn $\tau = (a, b)$ in $T_m(f, G)$ such that, for instance, $E(\tau, m) = \{a\}$ and $w(\tau, f) = v_{m-1}$, then $\Delta(f, G) > 0$ and several cases have to be considered.

(1) *If* $\delta(\tau, f) = d(f, G) \neq 0, \infty$. In this case we claim that, after one folding operation at τ, we obtain a topological representative (f', G') such that $d(f', G') \leqslant d(f, G) - 1$.

The proof of this claim is similar to the proof of Lemma 2.2.5. Indeed if $d(f, G) = j \neq 0$ then the sequence of vertices $\{v_1, \dots, v_{m-1}\}$ associated to the tangent turn τ is such that $Val(v_j) = 3$. After the folding operation at τ, the sequence of turn $\{\tau_1, \dots, \tau_{m-1}\}$ is transformed into $\{\tau'_1, \dots, \tau'_{m-1}\}$ and the turn $\tau'_1 = (a'_1, b'_1)$ is now tangent under f'. Furthermore, the vertex v_j has been transformed into v'_j which is also of valency 3 and thus the distance $\delta(\tau'_1, f') = j - 1$. Therefore $d(f', G') \leqslant d(f, G) - 1$, completing the proof of the claim.

If $\delta(\tau, f) = d(f, G) = \infty$, then again several cases are possible:

(2) *If* $\delta(\tau, f) = d(f, G) = \infty$ *and* $\Delta(\tau, f) = D(f, G) = 1$. In this case we claim that the folding operation at τ defines a topological representative (f', G') such that $d(f', G') < \infty$. Therefore the complexity function is decreasing and the new topological representative satisfies condition (1) above.

Indeed, with the above notations this situation occurs if all the vertices of the sequence $\{v, v_1, \dots, v_m\}$ have valence larger than 3. The fact that $\Delta(\tau, f) = 1$ implies that $f(a) = M \cdot A$ and $f(b) = M \cdot B$, where A and B are non-trivial paths. Therefore the folding operation at τ is partial and thus creates a new vertex ω of valency 3. The tangent turn as above is τ'_1 and the corresponding sequence of vertices is now $\{v'_1, \dots, v'_{m-1}, \omega\}$ and thus $\delta(\tau'_1, f') < \infty$, which proves the claim.

(3) *If* $\delta(\tau, f) = d(f, G) = \infty$ *and* $1 < \Delta(\tau, f) = D(f, G) < \infty$. Here a similar argument (shifting in the sequence of turns) shows that after a folding operation at τ we obtain a topological representative (f', G') such that

$$D(f', G') \leqslant D(f, G) - 1.$$

The very last case, i.e. $D(f, G) = \infty$, is covered by Proposition 2.2.6. In this case the topological representative (f, G) admits an invariant forest and, after collapsing this forest, we get $e(G') < e(G)$. This completes the proof of the main theorem.

Let us now make some observations about this version of the algorithm, compared with the one of [2]. The first difference comes from the folding operation itself which is done in a single algebraic step and requires less manipulations. For instance no valence two vertices have been introduced which suppress the need of removing these vertices by some valence two isotopies. This makes a big difference from an algorithmic point of view since it suppresses the need of any numerical computation.

The second main difference is the introduction of the complexity function which enables one to measure, at each step, the progress of the algorithm. The last difference comes from the case $m(f, G) > 1$ in the proof of Theorem 2.2.8. In Bestvina–Handel's approach, this case is treated by performing first a folding operation at the place where the cancellation occurs for the $(m + 1)$st iterate of the map, i.e. at a valence two vertex. This folding introduces a valence one vertex and also an extra valence 3 vertex. These unnecessary vertices are removed, in a second step, by a sequence of valence one and valence two isotopies (the valence 3 vertex is transformed into a valence 2 vertex after the valence one isotopy). This is not necessary here because we first perform a folding operation at the tangent vertex whose effect is either to decrease the tangent coefficient or one of the distances m, δ or Δ.

After a sequence of such reductions we arrive at a situation where the next folding decreases the growth rate. Furthermore it is easy to get an upper bound, in terms of the invariants m, δ and Δ, for the number of operations which are needed before the growth rate goes down.

A last comment has to be made about this new version of the algorithm. For simplicity we have excluded all the periodic automorphisms. In fact there is only one very specific case which creates a problem. This case is the one (and the only one) which occurs during the proof of Lemma 2.2.6. This case requires a special treatment (another folding operation). We have avoided this case to simplify the presentation.

3. A TRIP THROUGH THE TRACKS

3.1. Some basics

Let R_n be the rose with n petals and let $\mathcal{G}_n$ be the set of all the marked graphs whose fundamental group is F_n. We have already defined two basic operations which transform an element of $\mathcal{G}_n$ to another element of $\mathcal{G}_n$, namely the collapsing and the splitting operations. From the work of Culler and Vogtmann [3] we have:

LEMMA 3.1.1. *For any pair of marked graphs G, G' in $\mathcal{G}_n$ there is a finite sequence of collapsing and splitting operations which transforms G to G'.*

The next lemma is a reformulation, which can be found in [3], of a classical result due to Whitehead [18].

LEMMA 3.1.2. *A sequence of splitting and collapsing operations which transforms a rose R_n to another rose ρ_n induces an automorphism $h: F_n \to F_n$, and the converse holds.*

Lemma 3.1.1 enables one to define a distance on $\mathcal{G}_n$.

Definition 3.1.3. Let $d: \mathcal{G}_n \times \mathcal{G}_n \to \mathbf{N}$ be such that, for any pair (G, G') of marked graphs, $d(G, G')$ is the minimum number of splitting and collapsing operations which are needed to transform G to G'.

The function d is clearly a distance on $\mathcal{G}_n$.

3.2. Evolution paths

In the space $\mathscr{G}_n$ of all the marked graphs, if a path is an ordered sequence of splitting and collapsing operations, then Lemma 3.1.1 is interpreted as a path connectedness property. The goal of this section is to define a notion of path in the space $\mathscr{R}_n(\Phi)$ of all the topological representatives of $\Phi \in Out(F_n)$ and then to prove that this space is path connected.

The algorithm of Section 2, when applied to any $(f, G) \in \mathscr{R}_n(\Phi)$, defines an ordered sequence $\{(f_i, G_i)\}_{i\in[1,N]}$ in $\mathscr{R}_n(\Phi)$ which is ordered by the complexity function $\Lambda(f_i, G_i)$. We can choose this sequence in such a way that $d(G_i, G_{i+1}) \leqslant 1$. Indeed it suffices to decompose the collapsing of a forest into a sequence of elementary collapsing operations, i.e. edge by edge.

The ordering induced by the complexity function is very important for the algorithm itself but is too rigid for our next goal, so we shall soften it a little bit.

An ordered sequence $\{(f_i, G_i)\}_{i\in[1,N]}$ in $\mathscr{R}_n(\Phi)$ is called a *direct forward evolution path* if each transformation $(f_i, G_i) \to (f_{i+1}, G_{i+1})$ is an elementary move of the algorithm of Section 2, i.e. it is either a folding operation or a collapsing of an invariant or a pretrivial forest.

If we reverse the order of a direct forward evolution path then we obtain an ordered sequence in $\mathscr{R}_n(\Phi)$ which we call a *direct backward evolution path.*

Notice that the folding operation by itself, as defined in Section 2 at a tangent turn τ, is well defined even if the tangent coefficient $c(\tau, f) = \infty$; in which case we shall call the operation a *quasi-folding.*

We call a *forward evolution path* an ordered sequence $\{(f_i, G_i)\}_{i\in[1,N]}$ in $\mathscr{R}_n(\Phi)$ so that each elementary transformation $(f_i, G_i) \to (f_{i+1}, G_{i+1})$ is either an elementary move of the algorithm of Section 2 or a quasi-folding operation. A *backward evolution path* is obtained from a forward evolution path by reversing the order.

Definition 3.2.1. An *evolution path* connecting two topological representatives (f, G) and (f', G') is an ordered sequence $\{(f_i, G_i);\ i \in [0, N]\}$ in $\mathscr{R}_n(\Phi)$ such that

(i) $(f_0, G_0) = (f, G)$ and $(f_N, G_N) = (f', G')$, and

(ii) each sub-path $\{(f_i, G_i), (f_{i+1}, G_{i+1})\}$ is either a forward or a backward evolution path.

From this definition it is obvious that reversing the order of an evolution path gives also an evolution path.

3.3. $\mathscr{R}_n(\Phi)$ is path connected

The goal of this section is to prove the following:

THEOREM 3.3.1. *Let (f, G) and (f', G') be two topological representatives in $\mathscr{R}_n(\Phi)$, then there exists a finite evolution path connecting (f, G) and (f', G').*

In other words, with the notion of path given by the evolution paths, the space $\mathscr{R}_n(\Phi)$ is path connected. The idea to prove this theorem is to find an evolution path connecting any topological representative (f, G) of Φ to a known topological representative (base point). The natural choice of a particular representative is the initial map $\varphi: R_n \to R_n$ on a rose.

Lemma 3.3.2. *Let (f, G) be in $\mathscr{R}_n(\Phi)$ and let (φ, R_n) be a topological representative of Φ on a rose R_n. Then there exists a vertex $v \in V(G)$ and a maximal tree T in G such that collapsing T on the vertex v transforms (f, G) to (φ, R_n).*

By definition, for a given vertex $v \in V(G)$, there is an identification of $\pi_1(G, v)$ with F_n. Any topological representative $f: G \to G$ induces an automorphism $f_\#: \pi_1(G, v) \to \pi_1(G, v)$ of the fundamental group which differs from Φ by an inner automorphism. On the other hand $\varphi: R_n \to R_n$ induces $\varphi_\#: \pi_1(R_n, v_0) \to \pi_1(R_n, v_0)$ which differs from $f_\#$ by an inner automorphism. By choosing another vertex and perhaps another identification we can assume that $f_\# = \varphi_\#$.

Choosing an identification and a vertex is the same as choosing a maximal tree T in G, a vertex $v \in V(G)$ and a labelling of the edges of $G - T$. Once this choice is made, we obtain (φ, R_n) out of (f, G) by collapsing the maximal tree T on the vertex v.

Since G is finite, any maximal tree in G is finite and therefore Lemma 3.3.2 reduces the problem of finding an evolution path in $\mathscr{R}_n(\Phi)$ to the study of finitely many collapsing operations.

Proposition 3.3.3. *Let (f, G) be a topological representative of $\Phi \in Out(F_n)$. The collapsing of an edge $\{e\}$ in G such that $i(e) \neq t(e)$ defines (f', G') in $\mathscr{R}_n(\Phi)$ such that the transformation $(f, G) \to (f', G')$ can be decomposed into a finite evolution path.*

The proof of this proposition is purely combinatorial. Several preliminary observations are necessary before starting the proof of the proposition:

(1) If $G \in \mathscr{G}_n$, the collapsing of an edge $e \in E(G)$ defines a graph $G' \in \mathscr{G}_n$ if and only if $i(e) \neq t(e)$.
(2) The collapsing operation of an edge $e \in E(G)$, for a topological representative $(f, G) \in \mathscr{R}_n(\Phi)$, can be decomposed into two steps:
   (i) collapsing the edge path $f(e)$,
   (ii) collapsing the edge $\{e\}$.

This second step (ii) is actually a collapsing of a pretrivial forest and is therefore a forward evolution path.

If the edge $\{e\}$ belongs to an invariant forest then the collapsing of $\{e\}$ and its image $f(e)$ is a forward evolution path. We shall therefore assume, in the remaining of this section, that $\{e\}$ does not belong to an invariant forest. In this case we observe further that the collapsing of an edge path $f(e)$ is an oriented operation, i.e. it is defined from one extreme vertex to the other. If we change this orientation then we get, in general, another topological representative.

Lemma 3.3.4. *Let (f, G) be a topological representative of $\Phi \in Out(F_n)$ and let (f', G') be obtained from (f, G) by collapsing the edge path $f(e)$. If the collapsing is performed from $t(f(e))$ to $i(f(e))$ and such that all the turns at $v = t(e)$ which contain the edge $\{e\}$ are transverse under f, then the collapsing operation can be decomposed into an evolution path.*

At the vertex v we assume that $St(v) = \{e, e_1, \dots, e_{k-1}\}$ and we denote the edge path $f(e)$ by M. For each edge $x \in E(G) - St(v)$ the collapsing operation (of the edge $\{e\}$) is obtained by removing all the occurrences of the letter e (resp. $\bar{e}$) from the edge path $f(x)$. This is a rewriting operation which we denote by $f(x) \to (f(x))^{(*e)}$.

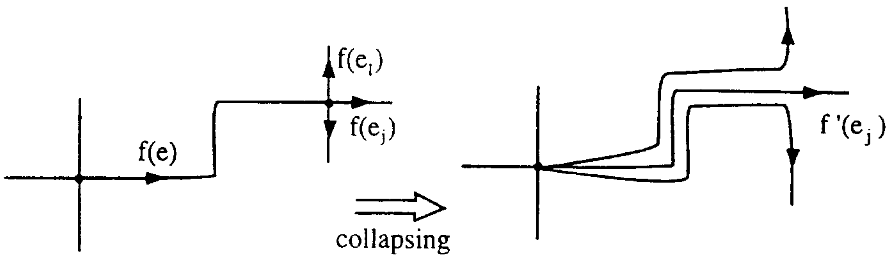

Fig. 5. Collapsing with transverse turns.

Let us now define the images of the other edges of G', i.e. $\{e_1, \ldots, e_{k-1}\}$ under the new map f'. In order to simplify the notations, we assume, after changing the orientations if necessary, that $v = i(e_j)$ for all $j = 1, \ldots, k-1$. With this orientation we have

$$f'(e_j) = M^{(*e)} \cdot f(e_j)^{(*e)} \quad \text{(see Fig. 5).} \qquad (\#)$$

Since all the turns which contain the edge $\{e\}$ at the vertex v are transverse under f, the new map is a topological representative. Indeed no cancellations are possible with the new map f' as defined by the above edge paths.

Let us now consider the following cases:

(a) If $Val(v) = k = 3$. In this case we have $St(v) = \{e, e_1, e_2\}$ and the edge $\{e\}$ connects v to $w = i(e)$ such that $St(w) = \{e, b_1, b_2, \ldots, b_{r-1}\}$. The collapsing operation identifies v with w and we denote by w' the resulting vertex of G'. This vertex is such that $St(w') = \{e_1, e_2, b_1, b_2, \ldots, b_{r-1}\}$, after an obvious identification of $E(G) - \{e\}$ with $E(G')$. The collapsing operation creates a tangency at the turn $\tau = (e_1, e_2)$ of w' due to the transformation $(\#)$. Indeed $M^{(*e)} \neq \emptyset$ since $f(e) \neq e$ because otherwise $\{e\}$ would be an invariant tree.

Now we observe that the transformation $(f', G') \to (f, G)$ is either a folding or a quasi-folding operation, depending on the tangent coefficient at τ. Therefore the ordered sequence $\{(f, G); (f', G')\}$ is an evolution path which completes the argument in this case.

(b) If $Val(v) = k > 3$. In this case the idea is to decompose the collapsing operation into several splitting–collapsing operations (see Fig. 6). In order to fix the notation we assume that

$$St(v) = \{e, e_1, e_2, \ldots, e_{k-1}\} \quad \text{and} \quad St(w) = \{e, b_1, b_2, \ldots, b_{r-1}\}.$$

After the collapsing operation, these two vertices are identified and the resulting vertex w' of G' is such that $St(w') = \{e_1, e_2, \ldots, e_{k-1}, b_1, b_2, \ldots, b_{r-1}\}$.

Step 1: *splitting*. As an intermediate step let us assume that the length of the edge path satisfies $|f(e)| = |M| \geqslant 2$ and we write $M = M_1 \cdot M_2$. We consider a new graph G_1 obtained from G by a splitting operation at the vertex v, i.e. with a new edge $\{x\}$ connecting the vertices v_1 and v_2 such that $St(v_2) = \{x, e_2, \ldots, e_{k-1}\}$ and $St(v_1) = \{e, e_1, x\}$.

We also define a new map $f_1 : G_1 \to G_1$ obtained from (f, G) by the following rewriting operation for an edge path p in G:

(Σa) Each occurrence, in p, of the turn $(\bar{e}, e_1)$ (i.e. a subword $e \cdot e_1$ or $\bar{e}_1 \cdot \bar{e}$) is left unchanged.

(Σb) Each occurrence, in p, of the letter e (resp. $\bar{e}$), except those arising from (Σa), is transformed according to $e \to e \cdot x$ (resp. $\bar{e} \to \bar{x} \cdot \bar{e}$).

(Σc) Each turn (e_j, e_1) is transformed according to $\bar{e}_j \cdot e_1 \to \bar{e}_j \cdot \bar{x} \cdot e_1$ (resp. $\bar{e}_1 \cdot e_j \to \bar{e}_1 \cdot x \cdot e_j$).

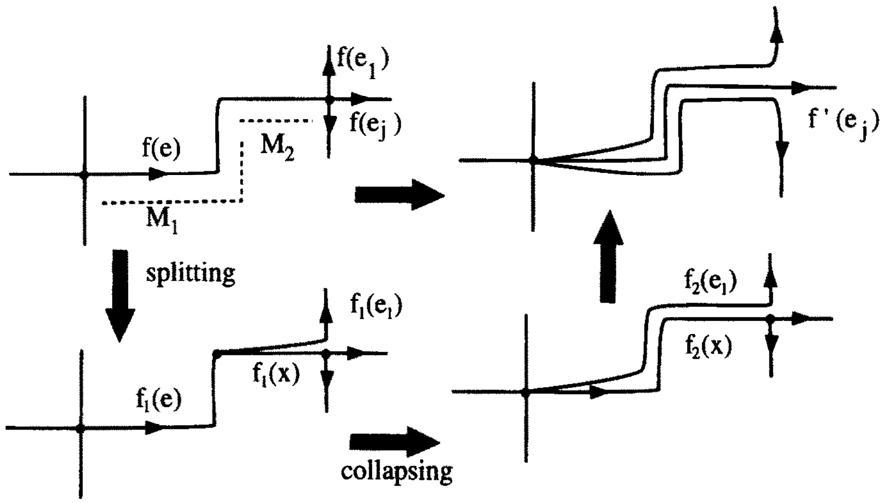

Fig. 6. Splitting–collapsing.

This transformation, denoted by $\Sigma(p)$, is a rewriting of an edge path in G to an edge path in G_1. The map f_1 is now defined by

$$f_1(y) = \Sigma(f(y)) \quad \text{for } y \in E(G_1) - \{e, e_1, x\}$$

and

$$f_1(x) = \Sigma(M_2), \;\; f_1(e) = \Sigma(M_1), \;\; f_1(e_1) = \Sigma(M_2)\cdot\Sigma(f(e_1)).$$

Again this new map is a topological representative of Φ, by the arguments used in part (a). As above, depending on the tangent coefficient $c((e_1, x); f_1)$, the transformation $(f_1, G_1) \to (f, G)$ is either a folding or a quasi-folding operation. Therefore the ordered pair $\{(f, G); (f_1, G_1)\}$ is an evolution path.

Step 2: *collapsing*. In order to complete the collapsing operation we have now to collapse the new edge path $f_1(e)$, from $f_1(v_1)$ to $f_1(w)$. But $Val(v_1) = 3$ and the turns (e, e_1), (e, x) are not tangent under f_1 by construction. Therefore the argument of part (a) applies for this new collapsing. We obtain a new topological representative (f_2, G_2) such that the collapsing operation $(f_1, G_1) \to (f_2, G_2)$ has been decomposed into a finite evolution path.

Let us observe that the last collapsing operation has identified the vertex v_1 with w. We denote by w_2 the resulting vertex and, with an obvious identification of $E(G_1) - \{e\}$ with $E(G_2)$, we have $St(w_2) = \{x, e_1, b_1, \dots, b_{r-1}\}$ while the vertex v_2 has not been changed, i.e. $St(v_2) = \{x, e_2, \dots, e_{k-1}\}$.

Now we want to iterate the previous splitting–collapsing operations. To that end we first have to collapse the edge path $f_2(x)$, from the vertex $f_2(v_2)$ to the vertex $f_2(w_2)$. But now the vertex v_2 has valence $k - 1$ and the steps 1, 2 above can again be applied. After $k - 3$ splitting–collapsing operations, we get a topological representative which is identified in an obvious way with (f', G'). Each splitting–collapsing operation is decomposed into a finite evolution path and therefore the transformation $(f, G) \to (f', G')$ is decomposed into a finite evolution path.

We assumed, at the begining of the proof, that $|f(e)| \geqslant 2$; this assumption has been used to simplify the writing of the map f_1 and to make clear what the splitting operation was. This assumption is a useful intermediate step but is not necessary. Indeed the splitting operation is followed by a collapsing operation. If $|f(e)| = 1$ then we can choose $M_1 = \emptyset$ in the definition of the map f_1. Then the edge $\{e\}$ is a pretrivial forest for f_1 which can be collapsed in a second step. This completes the proof of Lemma 3.3.4.

The proof of Proposition 3.3.3 is then reduced to the case where at least one turn, say (e, e_1), is tangent under f. In this case we first assume that the turn (e, e_1) is the only turn containing the edge $\{e\}$ which is tangent under f at the vertex v. We write

$$f(e) = N \cdot A, \quad f(e_1) = N \cdot B.$$

Notice that we have changed the orientation of the edge $\{e\}$ (with respect to the previous lemma) in order to simplify the notations.

The following cases have to be distinguished:

(1) $A = \emptyset$, (2) $A \neq \emptyset$ and $B \neq \emptyset$, (3) $B = \emptyset$.

Case (1): If $Val(v) = 3$ then the collapsing of the edge path $f(e)$ is an absorbing folding and therefore the ordered pair $\{(f, G); (f', G')\}$ is a forward evolution path. If $Val(v) = k > 3$ then we first perform an absorbing folding which removes the tangent turn (e, e_1). This folding defines a topological representative (f_1, G_1) (eventually after pulling tight) and the pair $\{(f, G); (f_1, G_1)\}$ is a forward evolution path. The graph G_1 has a new edge $\{x\}$ and a new vertex v' of valence $k - 1$. The collapsing of the edge path $f_1(x)$ can be decomposed into an evolution path by Lemma 3.3.4. Indeed the map f_1 has no tangency at the vertex v' (at a turn which contains the edge $\{x\}$). It gives rise to a topological representative which is identified with (f', G'), proving the proposition in this case.

Case (2): Assume again that $Val(v) = 3$. As a first step there is a partial folding which defines a forward evolution path. After this folding operation, the tangency has been removed and Lemma 3.3.4 can be applied in order to complete the collapsing of the edge path $f(e)$.

If $Val(v) = k > 3$ then we also apply first a partial folding operation which defines a topological representative (f_1, G_1) and the pair $\{(f, G); (f_1, G_1)\}$ is a forward evolution path. The graph G_1 has a new edge $\{x\}$ which connects a vertex v' of valence $k - 1$ to a new vertex v'' of valence 3. The collapsing of the two edge paths $f_1(x)$ and $f_1(e)$ can be decomposed into an evolution path by applying Lemma 3.3.4 twice. Indeed for these two edge paths there is no tangency at the vertex v' which contains the edge $\{e\}$ (because of the initial assumption) and also no tangency at v'' which contains the edge $\{x\}$, because the previous folding is partial.

Case (3): This is surely the most difficult case. We first observe that the collapsing of the edge path $f(e)$ from the topological representative (f, G) defines a map $\tilde{f}: \tilde{G} \to \tilde{G}$ which is not a topological representative since a cancellation occurs for the edge path $\tilde{f}(e_1)$ (see Fig. 7). After pulling tight we obtain a topological representative (f', G') and we want to prove that the transformation $(f, G) \to (f', G')$ can be decomposed into a finite evolution path.

In this case there are three important vertices in G, namely $v = i(e) = i(e_1)$, $w = t(e_1)$ and $u = t(e)$. In order to fix the notations we assume that $St(v) = \{e, e_1, \dots, e_{k-1}\}$, $St(w) = \{e_1, b_1, \dots, b_{r-1}\}$ and $St(u) = \{e, c_1, \dots, c_{q-1}\}$. The collapsing of the edge path $f(e)$ followed by the collapsing of $\{e\}$ identifies the vertices u and v. Let us call v' the resulting vertex of G'; it is such that $St(v') = \{e_1, \dots, e_{k-1}, c_1, \dots, c_{q-1}\}$.

Step 1: As in the previous cases we first apply a folding operation (absorbing) at $\tau = (e, e_1)$ in G. This operation defines a topological representative (f_1, G_1) (eventually after pulling tight). As above the pair $\{(f, G); (f_1, G_1)\}$ is a forward evolution path. The graph G_1 has a new edge $\{x\}$ which connects the vertex v_1 with the vertex w_1 such that

$$St(v_1) = \{x, e_2, \dots, e_{k-1}\} \text{ and } St(w_1) = \{x, e, b_1, \dots, b_{r-1}\}.$$

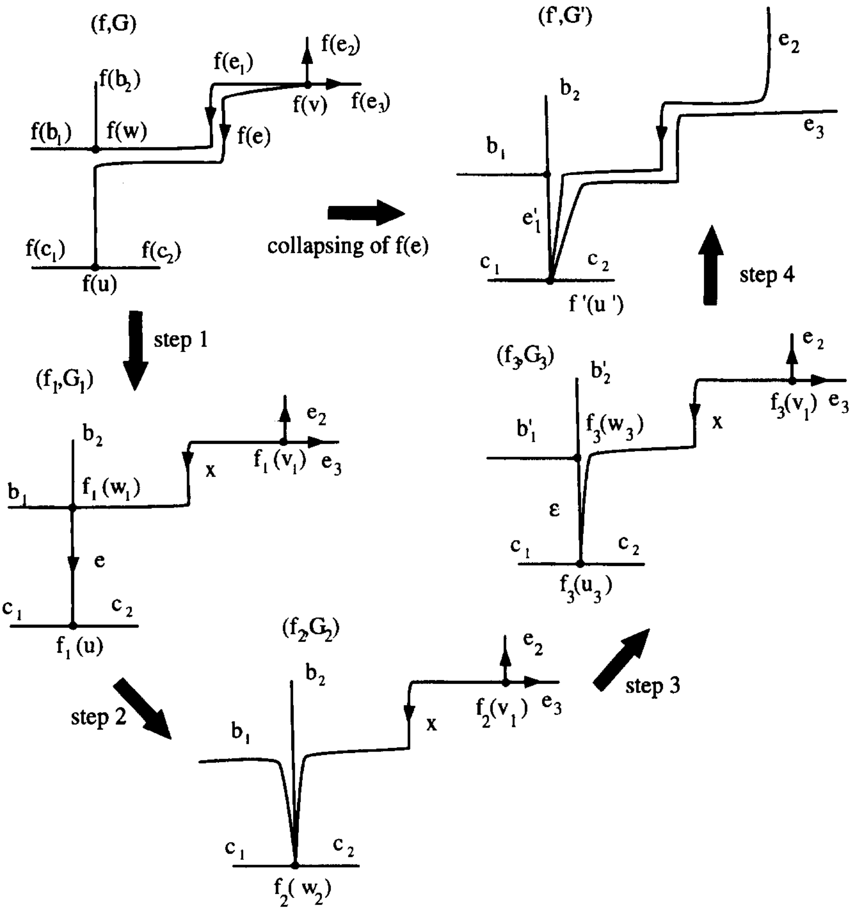

Fig. 7. Collapsing and evolution path.

For the topological representative (f_1, G_1) there are again two cases at the vertex w_1:

(i) No turn containing the edge $\{e\}$ is tangent under f_1 (see Fig. 7).
(ii) A new tangency has been created.

Let us first complete the proof in case (i).

Step 2: In case (i) we can apply Lemma 3.3.4 for the collapsing of the edge path $f_1(e)$. After this collapsing we obtain a topological representative (f_2, G_2) and, by Lemma 3.3.4, the transformation $(f_1, G_1) \to (f_2, G_2)$ can be decomposed into a finite evolution path. The collapsing of the edge $\{e\}$ and its image $f_1(e)$ identifies the vertices w_1 and u of the graph G_1. We denote by w_2 the resulting vertex of the graph G_2 which, after the natural identification of $E(G_1) - \{e\}$ with $E(G_2)$, is such that

$$St(w_2) = \{x, b_1, \dots, b_{r-1}, c_1, \dots, c_{q-1}\}.$$

For the new topological representative (f_2, G_2) all the edges $\{x, b_1, \dots, b_{r-1}\}$ are tangent under f_2 at the vertex w_2.

Step 3: In this part, the goal is to remove the tangencies which have been created at the vertex w_2 between the edges $\{b_1, \dots, b_{r-1}\}$, while keeping the edge path $f_2(x)$ unchanged. To that end we apply a sequence of folding. For instance we start this sequence by applying a partial folding for the topological representative (f_2, G_2) at the tangent turn (b_1, b_2). After this sequence of $(r-2)$ folding operations we obtain a topological representative (f_3, G_3) where all the tangencies between the b_i's have been removed. Since each folding operation defines an elementary forward evolution path, the transformation $(f_2, G_2) \to (f_3, G_3)$ is decomposed into a finite evolution path. The particular vertices of the graph G_3 are w_3 and u_3 such that

$$St(w_3) = \{\varepsilon, b'_1, b'_2, \dots, b'_{r-1}\} \quad \text{and} \quad St(u_3) = \{\varepsilon, x, c_1, \dots, c_{q-1}\}.$$

Step 4: In this last part we have to collapse the edge path $f_3(x)$, where the edge (x) connects the vertex u_3 to the vertex v_1 such that $St(v_1) = \{x, e_2, \dots, e_{k-1}\}$. Let us recall that we assumed, in the beginning of this proof, that for the initial topological representative (f, G) the turn (e, e_1) was the only tangency containing $\{e\}$ at the vertex v. Since the steps 1, 2, 3 have not affected the edges $\{e_2, \dots, e_{k-1}\}$ (except perhaps for the rewriting operations) and the edges $\{e, e_1\}$ have been replaced by the edge $\{x\}$ after the first folding operation, no turn (x, e_i) for $i = 2, \dots, k-1$ at v_1 is tangent under f_3. Therefore we can again apply Lemma 3.3.4 for the collapsing of the edge path $f_3(x)$ and we obtain finally a topological representative (f_4, G_4) such that the transformation $(f_3, G_3) \to (f_4, G_4)$ can be decomposed into a finite evolution path.

It remains to check that (f_4, G_4) can be identified with (f', G'). The collapsing of the edge $\{x\}$ has identified the vertices u_3 and v_1. If we call v_4 the resulting vertex of G_4 then it satisfies $St(v_4) = \{\varepsilon, c_1, \dots, c_{q-1}, e_2, \dots, e_{k-1}\}$. Furthermore the vertex w_3 has not been affected by the collapsing of $\{x\}$ and then $St(w_3) = \{\varepsilon, b'_1, b'_2, \dots, b'_{r-1}\}$. In order to identify (f_4, G_4) and (f', G') we have to identify e_1 with ε and b'_i with b_i for all $i = 1, \dots, r-1$. After this identification, the graphs are the same and we check that the images of each edge are the same. This completes the proof of Proposition 3.3.3 in the case (i) defined at step 1.

In case (ii), i.e. when the topological representative (f_1, G_1) admits a tangency at a turn (e, b_i) of the vertex w_1, then again the cases (1), (2) and (3) above are possible. Therefore we have to iterate the previous arguments for collapsing the edge path $f_1(e)$. This iteration process must stop after finitely many steps since the length of the edge path $f_1(e)$ is strictly less than the length of the original edge path $f(e)$ and all these lengths are finite.

We assumed, for simplicity, that the initial topological representative (f, G) has only one tangent turn at v containing the edge $\{e\}$. If this assumption is not satisfied then we apply the same arguments for each such tangency. This completes the proof of Proposition 3.3.3. The steps for the proof of the above case 3 can be visualized by Fig. 7.

Proof of Theorem 3.3.1. Let (f, G) and (f', G') be two topological representatives of $\Phi \in Out(F_n)$. By Lemma 3.3.2 we can find a finite sequence of collapsing operations which transforms (f, G) to (φ, R_n) and another sequence which transforms (f', G') to (φ, R_n). By Proposition 3.3.3 each elementary collapsing operation can be decomposed into a finite evolution path. Therefore there is a finite evolution path which connects (f, G) to (φ, R_n) and another one which connects (f', G') to (φ, R_n). Since by reversing the order of an evolution path we get an evolution path, we have found a finite evolution path connecting (f, G) to (f', G').

3.4. Peak reduction

In this subsection the goal is to study the evolution of the growth rate along an evolution path in $\mathcal{R}_n(\Phi)$. We already observed that the growth rate is decreasing along a forward evolution path. The goal is to prove the following:

THEOREM 3.4.1 (Global peak reduction). *For every pair* (f, G) *and* (f', G') *of topological representatives of* $\Phi \in Out(F_n)$, *there is an evolution path* $\{(f_t, G_t)\}_{t\in[0,N]}$ *such that*

(i) $(f_0, G_0) = (f, G)$ *and* $(f_N, G_N) = (f', G')$,
(ii) $\lambda(f_t, G_t) \leqslant \max\{\lambda(f_0, G_0), \lambda(f_N, G_N)\}$, $\forall t \in [0, N]$.

This peak reduction result has an immediate corollary, when restricted to the set of train track representatives.

COROLLARY 3.4.2 (Path connectedness). *For every pair* (f, G) *and* (f', G') *of train track representatives of* $\Phi \in Out(F_n)$, *there is an evolution path* $\{(f_t, G_t)\}_{t\in[0,N]}$ *such that*

(i) $(f_0, G_0) = (f, G)$ *and* $(f_N, G_N) = (f', G')$,
(ii) (f_t, G_t) *is a train track representative for each* $t \in [0, N]$.

The proof of the corollary from the theorem is straightforward since any train track representative has the minimal growth rate among all the possible topological representatives. Conversely, any topological representative whose growth rate is minimal is a train track map.

Definition 3.4.3. A *local peak* in an evolution path $\{(f_t, G_t)\}_{t\in[0,N]}$ is a subpath containing three consecutive representatives $(f_{i-1}, G_{i-1}), (f_i, G_i), (f_{i+1}, G_{i+1})$ such that

$$\lambda(f_{i-1}, G_{i-1}) \leqslant \lambda(f_i, G_i) \geqslant \lambda(f_{i+1}, G_{i+1}).$$

The representative (f_i, G_i) is called a *local maximum*. If both inequalities are strict then the local peak is called a *strict local peak*.

Similarly we call the subpath $\{(f_t, G_t)\}_{t\in[j,j+k]}$ a *flat peak* if

$$\lambda(f_j, G_j) \leqslant \lambda(f_{j+1}, G_{j+1}) = \lambda(f_{j+2}, G_{j+2}) = \cdots = \lambda(f_{j+k-1}, G_{j+k-1}) \geqslant \lambda(f_{j+k}, G_{j+k}).$$

The proof of Theorem 3.4.1 is an analysis of the local and the flat peaks, the goal being to "erase" all the peaks in an evolution path.

LEMMA 3.4.4 (Local peak reduction). *If* $\{(f_t, G_t)\}_{t\in[0,2]}$ *is a local peak then there exists a finite evolution path* $\{(f_t', G_t')\}_{t\in[0,N]}$ *connecting* (f_0, G_0) *to* (f_2, G_2) *such that*

$$\lambda(f_t', G_t') \leqslant \max\{\lambda(f_0, G_0), \lambda(f_2, G_2)\}, \quad \forall t \in [0, N].$$

If the local peak is not strict then the lemma is obvious. At a local maximum (f_1, G_1) in a strict local peak there are two elementary forward evolution paths, one leading to (f_0, G_0), the other leading to (f_2, G_2). These elementary paths are necessarily two folding operations, since λ is strictly decreasing. Several cases are possible.

(a) *The two folding operations are "independent"*: A folding operation at a topological representative (f, G) is characterized by a pair of edges (a, b) which are incident at the same vertex $v \in V(G)$ and by a common edge path of their f-image. We say that two folding operations are *independent* if either the two pairs of edges are disjoint or, if one edge belongs to the two pairs, the two common edge paths are disjoint. In this case we can perform the two folding operations independently. After one folding from (f_1, G_1), one gets either (f_0, G_0) or (f_2, G_2). Each one of these topological representatives admits another folding which gives rise to

$$(f_0, G_0) \to (f_3, G_3) \quad \text{and} \quad (f_2, G_2) \to (f_4, G_4).$$

Let us now check that the two topological representatives (f_3, G_3) and (f_4, G_4) can be identified. In order to fix the notations, let us call (a, b) and (c, d) the two turns which are tangent under f_1. In this case we assume that

$$f_1(a) = M \cdot A, \quad f_1(b) = M \cdot B \quad \text{and} \quad f_1(c) = N \cdot C, \; f_1(d) = N \cdot D.$$

Let us call $\{x\}$ (resp. $\{y\}$) the new edge (if any) which is created by the folding operation $(f_1, G_1) \to (f_0, G_0)$ (resp. $(f_1, G_1) \to (f_2, G_2)$). If these two folding operations are partial then we identify $E(G_1) \cup \{x\}$ with $E(G_0)$ and $E(G_1) \cup \{y\}$ with $E(G_2)$. If one of these two folding operations is absorbing or if one of these two foldings occurs at a valence 3 vertex then the adaptations are obvious.

Now if we denote by $\{y\}$ (resp. $\{x\}$) the edge which is created by the folding operation $(f_0, G_0) \to (f_3, G_3)$ (resp. $(f_2, G_2) \to (f_4, G_4)$) then the two graphs G_3 and G_4 are the same and with the same labelling of the edges. For each edge $e_i \in E(G_3) - \{x, y, a, b, c, d\}$ then the edge paths $f_3(e_i)$ and $f_4(e_i)$ are obtained from the edge path $f_1(e_i)$ by the two successive rewriting operations:

$$(*x) \begin{cases} a \to x \cdot a \\ b \to x \cdot a \end{cases} \quad \text{and} \quad (*y) \begin{cases} c \to y \cdot c \\ d \to y \cdot d. \end{cases}$$

These two rewriting operations commute and therefore $f_3(e_i) = f_4(e_i)$.

By assumption the edge paths M and N in G_1 are disjoint and we have

$$f_3(x) = ((M)^{(*x)})^{(*y)}, \quad \text{while} \quad f_4(x) = ((M)^{(*y)})^{(*x)}.$$

The fact that the two rewriting operations commute imply that $f_3(x) = f_4(x)$. By the same arguments we prove that $f_3(e) = f_4(e)$ for all $e \in \{a, b, c, d\}$. Therefore $(f_3, G_3) = (f_4, G_4)$, and we have found an evolution path satisfying the properties of Lemma 3.4.4, namely $(f_0, G_0) \to (f_3, G_3) \to (f_2, G_2)$ (see Fig. 8(a)).

(b) *The two folding operations are not independent*: Again two cases are possible.

(i) *The two turns belong to the same vertex*: This case occurs if three edges (a, b, c) are incident at $v \in V(G_1)$ and, if we chose the orientation so that $v = i(a) = i(b) = i(c)$, then we have

$$f_1(a) = M \cdot A, \quad f_1(b) = M \cdot B \quad \text{and} \quad f_1(c) = M.C$$

where M is the longest initial common edge path between $f_1(a), f_1(b), f_1(c)$. Let us assume that the two turns in question are (a, b) and (b, c). Let us denote by P the longest initial edge path between $f_1(a)$ and $f_1(b)$ and by Q the longest initial edge path between $f_1(b)$ and $f_1(c)$. After one folding from (f_1, G_1) one gets either (f_0, G_0) or (f_2, G_2). Each one of these topological representatives admits another folding which gives rise to $(f_0, G_0) \to (f_3, G_3)$ and $(f_2, G_2) \to (f_4, G_4)$. At that point two cases are again possible:

(α) $|P| = |Q| = |M|$. In this case, as in case (a) above we check that we can identify (f_3, G_3) and (f_4, G_4); therefore the argument of part (a) applies.

(β) $|Q| = |M| < |P|$ (*the other case is similar*). In this case the two topological representatives (f_3, G_3) and (f_4, G_4) are different. Indeed the evolution path $(f_1, G_1) \to (f_0, G_0) \to (f_3, G_3)$ is obtained by two partial folding operations whereas in the other path $(f_1, G_1) \to (f_2, G_2) \to (f_4, G_4)$, the second folding operation is absorbing. Thus the two graphs G_3 and G_4 are not the same. This last topological representative (f_4, G_4) admits another folding at the new turn (a', b'), which is the image of the turn (a, b) under the above sequence of foldings. By performing this folding operation one gets another topological representative (f_5, G_5) and one checks that (f_3, G_3) and (f_5, G_5) can be identified.

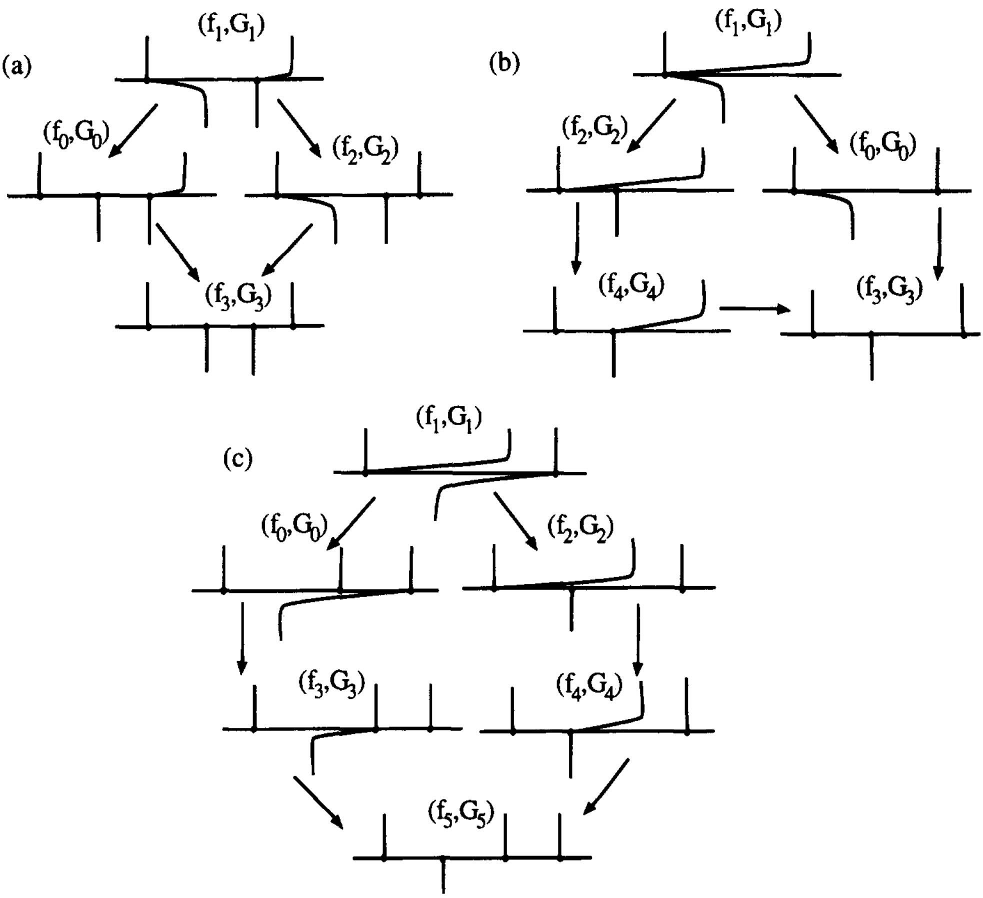

Fig. 8.

Therefore we have found an evolution path satisfying the properties of Lemma 3.4.4, namely

$$(f_0, G_0) \to (f_3, G_3) = (f_5, G_5) \to (f_4, G_4) \to (f_2, G_2) \quad \text{(see Fig. 8(b))}.$$

(ii) *The two turns belong to distinct vertices*: This case occurs if an edge (say $\{b\}$) connects two vertices v and w. We chose the orientation so that $v = i(a) = i(b)$ and $w = t(b) = i(c)$ and the turns $(a, b), (c, \bar{b})$ are tangent under f_1. We denote by M the longest initial edge path between $f_1(b)$ and $f_1(a)$ and by N the longest initial edge path between $f_1(\bar{b})$ and $f_1(c)$. By assumption the edge paths M and N are not disjoint. As in the above cases from (f_1, G_1) there are two different folding operations leading to (f_0, G_0) or (f_2, G_2). Again these two topological representatives admit another folding $(f_0, G_0) \to (f_3, G_3)$ and $(f_2, G_2) \to (f_4, G_4)$. Again, two cases are possible:

(α) *The intersection of the two edge paths M and N is restricted to a vertex.* In this case, we check as above that (f_3, G_3) and (f_4, G_4) can be identified. Therefore the arguments of part (a) apply.

(β) *The intersection of M and N is a non-trivial edge path.* In this case $(f_3, G_3) \neq (f_4, G_4)$ and each of these two topological representatives admits a folding. If we perform these two folding operations then one gets

$$(f_3, G_3) \to (f_5, G_5) \quad \text{and} \quad (f_4, G_4) \to (f_6, G_6).$$

As above we check that $(f_5, G_5) = (f_6, G_6)$, after suitable identification of the two graphs. Again we have found an evolution path satisfying the properties of Lemma 3.4.4, namely (see Fig. 8(c))

$$(f_0, G_0) \to (f_3, G_3) \to (f_5, G_5) = (f_6, G_6) \to (f_4, G_4) \to (f_2, G_2).$$

This last case completes the proof of the lemma.

LEMMA 3.4.5 (Flat peak reduction). *If $\{(f_t, G_t)\}_{t\in[0,k]}$ is a flat peak then there exists an evolution path $\{(f_t', G_t')\}_{t\in[0,N]}$ connecting (f_0, G_0) to (f_k, G_k) such that*

$$\lambda(f_t', G_t') \leqslant \max\{\lambda(f_0, G_0), \lambda(f_k, G_k)\}, \quad \forall t \in [0, N].$$

As in the previous lemma, if the peak is not strict (non-strict inequalities in the definition) then the lemma is obvious. As in the previous proof, there are two elementary forward evolution paths $(f_1, G_1) \to (f_0, G_0)$ and $(f_{k-1}, G_{k-1}) \to (f_k, G_k)$. Furthermore the topological representatives (f_1, G_1) and (f_{k-1}, G_{k-1}) are connected by a "flat" evolution path, i.e. whose topological representatives all have the same growth rate. There are three basic flat evolution paths:

(1) Collapsing an invariant forest.

(2) A sequence of folding operations such as the one occurring in the proof of Theorem 2.2.8.

(3) A sequence of quasi-folding operations.

Let us first assume that the flat part of the flat peak consists only of one of the above types.

Type (1): The collapsing of an invariant forest in the above flat peak is given by the sequence $(f_1, G_1) \to (f_2, G_2) \to \cdots \to (f_{k-1}, G_{k-1})$. At (f_{k-1}, G_{k-1}) there is a folding operation giving rise to (f_k, G_k) and at (f_1, G_1) there is another folding operation giving rise to (f_0, G_0). Let us focus our attention on the last folding of this sequence. We denote by (a, b) the turn which is tangent under f_{k-1}. These two edges $\{a\}$ and $\{b\}$ of G_{k-1} also belong to the graph G_1, after an obvious identification. Indeed the sets $E(G_1)$ and $E(G_{k-1})$ can be identified, except for the edges which have been collapsed, obviously the edges $\{a, b\}$ have not been collapsed. There are two possibilities for the graph G_1:

(a) *The edges $\{a\}$ and $\{b\}$ start at the same vertex $v \in V(G_1)$*: In this case the topological representative (f_1, G_1) admits two different folding operations. Indeed let us check that the turn (a, b) in G_1 is tangent under f_1. Assume that (a, b) is transverse under f_1. Then the following property is satisfied:

$$f_1(a) = X \cdot M \cdot A \quad \text{and} \quad f_1(b) = Y \cdot M \cdot B$$

where X and Y are two distinct edge paths in one connected component T of the f_1-invariant forest F. This situation cannot occur because otherwise the edge path $X \cdot \bar{Y}$ would be a non-trivial loop, which is impossible because T is a tree (see Fig. 9). Therefore the topological representative (f_1, G_1) admits two different folding operations $(f_1, G_1) \to (f_0, G_0)$ and $(f_1, G_1) \to (f_2', G_2')$ and also an invariant forest F. We thus have to prove that the collapsing of the invariant forest commutes with the folding operation $(f_1, G_1) \to (f_2', G_2')$.

In order to simplify the arguments we assume that the forest F is reduced to a single edge, i.e. $f_1(e) = e$.

(i) *If $e \notin St(v)$*. In this case the edge $\{e\}$ is not affected by the folding operation at the turn (a, b). Furthermore, since $f_1(e) = e$, the rewriting operation does not affect this edge path

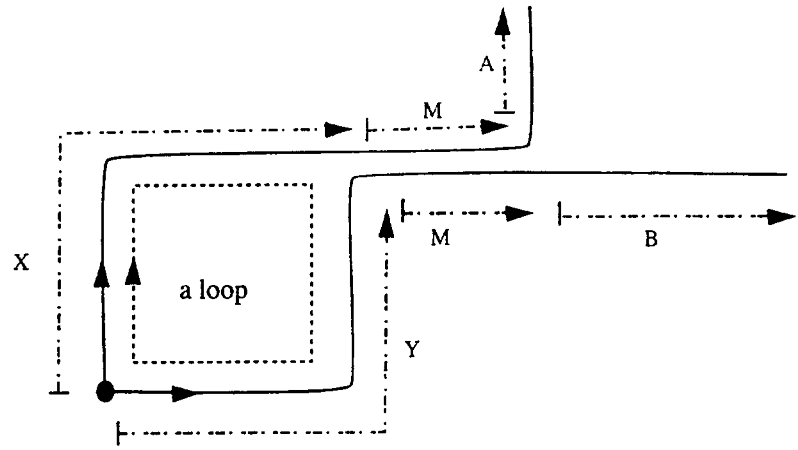

Fig. 9. No loop in a tree!

either (no occurrence of the letters a or b). Therefore the topological representative (f_2', G_2') has an invariant forest which is identified with F. The commutation of the folding operation and the collapsing operation is then clear in this case. If the tree is more complicated then the same arguments hold. Finally, in this case, the arguments of the proof of Lemma 3.4.4 can be applied directly here at the topological representative (f_1, G_1), i.e. we first apply the folding operation and then collapse the invariant forest. Finally we obtain an evolution path satisfying the properties of Lemma 3.4.5.

(ii) *If* $e \in St(v)$. If $Val(v) > 3$ then the previous arguments apply here to prove that the folding and the collapsing operation commute. Therefore the same conclusion holds. If $Val(v) = 3$ then the folding operation affects the existence of the invariant forest. Nevertheless we prove that the collapsing plus folding operations give the same topological representative as the other folding alone (in this case).

Indeed our assumptions imply that

$$f_1(a) = M.A, \quad f_1(b) = M.B \quad \text{and} \quad f_1(e) = e.$$

Then the folding operation gives (f_2', G_2') such that

$$f_2'(a) = A, \quad f_2'(b) = B \quad \text{and} \quad f_2'(e) = e.M.$$

Let us observe that, since $f_1(e) = e$, the two vertices $w = i(e)$ and $v = t(e)$ are fixed under f_1. This implies that the edge path M is either $M = a \cdot M'$ or $M = b \cdot M''$.

Let us now apply first the collapsing operation which gives

$$f_2(a) = (M \cdot A)^{(*e)} \quad \text{and} \quad f_2(b) = (M \cdot B)^{(*e)}$$

where $(*e)$ is the rewriting operation defined in Section 2.2 (1a). The turn (a, b) is now at the vertex w' which is obtained by the identification of v and w; this vertex is such that $Val(w') > 3$. The folding operation, at the turn (a, b), creates a new edge $\{x\}$ for a topological representative (f_3, G_3) which satisfies

$$f_3(a) = (A)^*, \quad f_3(b) = (B)^* \quad \text{and} \quad f_3(x) = (M)^*.$$

We observe that $E(G_3)$ is identified with $E(G_2')$ if we set $\{x\} = \{e\}$. With the above properties we check that $f_3 = f_2'$. The arguments are the same if the invariant forest is more complicated. Therefore we can again apply the Lemma 3.4.4 at the topological representative (f_1, G_1) which completes the proof of Lemma 3.4.5 in this case.

(b) $\{a\}$ *and* $\{b\}$ in G_1 *starts at different vertices*: $a \in St(v)$ *and* $b \in St(w)$: Since (a, b) is a turn in G_{k-1} then the vertices v and w in G_1 are connected by an edge path P which is contained in one component T of the f_1-invariant forest F. Furthermore the turn (a, b) is tangent under f_{k-1}, then the edge paths $f_1(a)$ and $f_1(b)$ can be written as

$$f_1(a) = T_1 \cdot M \cdot A \quad \text{and} \quad f_1(b) = T_2 \cdot M \cdot B$$

where T_1 and T_2 are two edge paths in the component $f_1(T)$ of F. As a first step we observe, since T is a tree, that the edge path $T_1 \cdot \bar{T}_2$ contains no trivial loops and then $T_1 \cdot \bar{T}_2$ is homotopic to $f_1(P)$. With the previous observations, there are two edges: $t_1 \in E(T) \cap St(v)$ and $t_2 \in E(T) \cap St(w)$, which are the extreme edges of the edge path P defined above. Furthermore the turns (a, t_1) and (b, t_2) are tangent under f_1 and thus we can apply a folding operation (absorbing) at these turns. After these two folding operations (which commute by the proof of Lemma 3.4.4) we obtain a topological representative (f_1', G_1'). We check that (f_1', G_1') admits an invariant forest F' which can be identified with F. The new edge path $f_1'(a)$ (resp. $f_1'(b)$) has an initial edge path T_1' (resp. T_2') in F' whose length is strictly less than T_1 (resp. T_2). This argument can be used inductively and we obtain a finite sequence of absorbing folding operations: $(f_1, G_1) \to (f_1', G_1') \to \cdots (f_j', G_j')$ such that:

— (f_j', G_j') has an invariant forest F_j' which can be identified with F.
— The two edges $\{a\}$ and $\{b\}$ form a turn (a, b) which is tangent under f_j'.

In other words the topological representative (f_j', G_j') satisfies the assumptions of case (a) above. Since the forest F_j' is identified with the forest F we can collapse it which gives a new topological representative (f_{j+1}', G_{j+1}'). We check that it can be identified with (f_{k-1}, G_{k-1}). This proves that the sequence of absorbing folding operations, $(f_1, G_1) \to (f_1', G_1') \to \cdots \to (f_j', G_j')$, commutes with the sequence of collapsing operations, $(f_1, G_1) \to (f_2, G_2) \to \cdots \to (f_{k-1}, G_{k-1})$. Finally the arguments of case (a) can be applied at the topological representative (f_j', G_j'), proving Lemma 3.4.5 in this type 1 case, by the same arguments.

Type (2): In this case we only have to deal with a sequence of folding operations. Let us observe that we have not used the fact, in the proof of Lemma 3.4.4, that the two folding operations decrease the growth rate. In this type (2) case, the analysis of the situation where, at a given topological representative (f_i, G_i) of the flat part sequence, there are two distinct folding operations is the same as for Lemma 3.4.4. Therefore the same conclusion holds, proving Lemma 3.4.5 in this case.

Type (3): In this case we assumed that the sequence $(f_1, G_1) \to (f_2, G_2) \to \cdots \to (f_{k-1}, G_{k-1})$ is induced by a sequence of quasi-folding operations. The previous observations holds which completes the proof of Lemma 3.4.5.

Proof of Theorem 3.4.1. From Theorem 3.3.1 we can find a finite evolution path $\{(f_t, G_t)\}_{t \in [0,N]}$ connecting any pair (f, G) and (f', G') of topological representatives of $\Phi \in Out(F_n)$. Since this path is finite, there is a *global maximum representative*, i.e. there exists (f_M, G_M) in $\{(f_t, G_t)\}_{t \in [0,N]}$ such that $\lambda(f_M, G_M) \geqslant \lambda(f_t, G_t)$ for all $t \in [0, N]$. If this global maximum is not strict in an obvious sense then we are done. If this maximum is strict then it defines either a local or a flat peak. By the previous Lemmas 3.4.4 and 3.4.5 we can "reduce" the peak in finding another evolution path with a lower maximum. In order to complete the proof we have to iterate this peak reduction process. Since at each step we get a finite path, the iteration is finite.

4. THE CONJUGACY PROBLEM

For a given outer automorphism Φ we denote by $\mathscr{T}(\Phi)$ the set of all the train track representatives of Φ. The first property is obvious:

$\mathscr{T}(\Phi)$ is finite.

Indeed all the train track representatives have the same growth rate λ which is the largest eigenvalue of an irreducible integer matrix of bounded size.

Definition 4.1. Two topological representatives (f_1, G_1) and (f_2, G_2) are *equal* if there is a one to one map $\alpha: G_1 \to G_2$ such that $f_2 = \alpha^{-1} \circ f_1 \circ \alpha$.

Notice that this notion of equality is the simplest possible one. Furthermore it is easy to check in practice. Indeed the existence of the map α means that the graphs have the same topological structure and α is just a renaming of the edges. After this renaming the edge paths of the f_1 and f_2 images are the same. The main result of this section is:

THEOREM 4.2. *Two irreducible outer automorphisms Φ_1 and Φ_2 are conjugated if and only if $\mathscr{T}(\Phi_1) = \mathscr{T}(\Phi_2)$.*

Notice that the path connections result 3.4.2 implies that $\mathscr{T}(\Phi_1) = \mathscr{T}(\Phi_2)$ if and only if $\mathscr{T}(\Phi_1) \cap \mathscr{T}(\Phi_2) \neq \emptyset$.

(a) *If Φ_1 and Φ_2 are conjugated.* Let us consider $\{x_1, \dots x_n\}$ a set of generators of the free group F_n. This set defines a marking of the rose R_n for the initial topological representative (φ_1, R_n). If we apply the algorithm of Section 2 to this initial topological representative we obtain, after finitely many steps, a train track representative (f_1, G_1). Since Φ_1 and Φ_2 are conjugated, there is an automorphism h such that $\Phi_2 = h \circ \Phi_1 \circ h^{-1}$. The set $\{h(x_1), \dots, h(x_n)\}$ is also a generating set for F_n and it defines another marking of the rose, which we denote by hR_n. The image of this generating set under Φ_2 is given by a collection of n words which are the same as the image of the generating set $\{x_1, \dots, x_n\}$ under ϕ_1, since Φ_1 and Φ_2 are conjugated by h. With this particular initial marking, if we apply the algorithm of Section 2 to (φ_2, R_n) one gets a train track representative for Φ_2. Since the words are the same for (φ_1, R_n) and for (φ_2, hR_n), we can follow the same evolution path leading to the train track (f_1, G_1); therefore $\mathscr{T}(\Phi_1) \cap \mathscr{T}(\Phi_2) \neq \emptyset$.

(b) $\mathscr{T}(\Phi_1) \cap \mathscr{T}(\Phi_2) \neq \emptyset$. Under this assumption the two outer automorphisms have a common topological representative (f, G). This means that Φ_1 has a train track representative (f_1, G_1) and Φ_2 has a train track representative (f_2, G_2) which can be identified by $\alpha: G_1 \to G_2$. This map induces $\alpha_\#: \pi_1(G_1, v_1) \to \pi_1(G_2, v_2)$ which is an automorphism of the free group F_n. Therefore Φ_1 and Φ_2 are the same, up to a renaming of the generators of F_n defined by $\alpha_\#$. This means that Φ_1 and Φ_2 are conjugated.

The conjugacy problem: Let us make the solution to the conjugacy problem explicit. We start with two automorphisms Φ_1 and Φ_2 and we want to decide whether they are conjugated or not. To this end we apply the following procedure:

(1) Find a train track representative (f_1, G_1) for Φ_1 and (f_2, G_2) for Φ_2.

(2) If the growth rates $\lambda(f_1, G_1)$ and $\lambda(f_2, G_2)$ are different then the two automorphisms are not conjugated.

(3) If the two growth rates are the same then from the path connectedness property (Corollary 3.4.2) we want to find an evolution path connecting (f_1, G_1) to (f_2, G_2). If we can

find such a path then the two automorphisms are conjugated; if not, they are not conjugated.

The path connectedness property implies that we can restrict ourselves to the evolution paths all of whose topological representatives are train tracks. For such an evolution path, each elementary transformation is either a quasi-folding operation or a collapsing of an invariant forest.

Since the set $\mathcal{T}(\Phi_1)$ is finite we can generate the whole set $\mathcal{T}(\Phi_1)$ by a finite sequence of quasi-folding operations or collapsing invariant forests. In order to solve the conjugacy problem it just remains to check whether (f_2, G_2) belongs or not to the finite set $\mathcal{T}(\Phi_1)$ which we constructed above. This completes the proof of the main theorem.

Remarks—questions. (1) As we noticed in the introduction, Theorem 1 is not sufficient to solve completely the conjugacy problem in $Out(F_n)$. It remains to understand the reducible cases. From a graph point of view the reducibility is detected by the existence of a topological representative which admits an invariant sub-graph whose components have a non-trivial fundamental group.

What is a "good" notion of a train track in this case (an example of such train tracks is given in [2])? Is there a "natural" (for the conjugacy questions) sub-class? Is there a "natural" notion of Dehn twist, relative to a subgraph, as in the surface case?

(2) For practical questions it would be useful to find a complexity estimate for the solution described above. In particular, for part (3) above, an estimate of the cardinality of $\mathcal{T}(\Phi)$ would be necessary.

(3) Theorem 2.2.8 indicates that the algorithm, which starts at a given topological representative and stops at a train track map, could be interpreted as a dynamical system in $\mathcal{R}_n(\Phi)$. This dynamics looks like a gradient flow, i.e. with a function which is decreasing along the orbits (forward evolution paths). We have used this analogy in [12] to define a variational calculus in $\mathcal{R}_n(\Phi)$ for which the train track representatives are critical points of a certain functional. The path connectedness property can be interpreted in this setting. This analogy is promising and could possibly help to analyse other classes of algorithms.

REFERENCES

1. M. Bestvina and M. Handel: Train tracks for surface automorphisms, *Topology* **34** 109–140 (1995).
2. M. Bestvina and M. Handel: Train tracks and automorphisms of free groups, *Ann. of Math.* **135** (1992), 1–51.
3. M Culler and K. Vogtmann: Moduli of graphs and automorphisms of free groups, *Invent. Math.* **84** (1986), 91–119.
4. M. Dehn: Uber unendliche diskontinnerliche gruppen, *Math. Ann.* **71** (1912), 116–144.
5. W. Dicks: Groups, trees and projective modules, Springer Lecture Notes **790** Springer, Berlin (1980).
6. W. Dicks and E. Ventura: Irreducible automorphisms of growth rate one, *J. Pure Appl. Algebra* **88** (1993), 51–62.
7. F. R. Gantmacher: *The theory of matrices*, Chelsea, New York (1959).
8. S.M. Gersten: Fixed points of automorphisms of free groups, *Adv. in Math.* **64** (1987), 51–85.
9. G. Hemion: On the classification of homeomorphisms of 2-manifolds and the classification of 3-manifolds, *Acta. Math.* **1–2** (1979), 125–155.
10. W. Hurwicz: A survey of the conjugacy problem, *Contemp. Math.* **33** (1984), 278–298.
11. J. Los: Pseudo-Anosov maps and invariant train tracks in the disc: a finite algorithm, *Proc. London Math. Soc.* (3) **66** (1993), 400–430.
12. J. Los: A variational calculus for automorphisms of free groups, in preparation.
13. M. Lustig: Automorphisms, train tracks and non-simplicial R-tree actions, Preprint, Bochum (1993).
14. L. Mosher: The classification of pseudo-Anosovs, *London Math. Soc. Lect. Notes* **112** (1986), 13–75.

15. J. NIELSEN: Die struktur periodischer transformationen von flachen, *Det. Kgl. Dansk. Uidensk. Selskab. Mat. fys. Meddelever* **15** (1937) 1–77: or in *Collected work*, Birkhauser, (1984).
16. J. R. STALLINGS: Topology of finite graphs, *Invent. Math.* **71** (1983), 551–565.
17. W. P. THURSTON: On the geometry and the dynamics of diffeomorphisms of surfaces, *Bull. AMS* **19** (1988), 415–431.
18. J. H. C. WHITEHEAD: On the equivalent sets of elements in a free group, *Ann. Math.* **37** (1936), 782–800.

Institut Non Linéaire de Nice
CNRS, UMR 129
Université de Nice-Sophia Antipolis
1361 Route des Lucioles
06 560 Valbonne
France